

## S6 Alignment. Alignment of the X protein of erythroparvoviruses (including reported endogenous erythroparvoviruses) with the ARF1 protein of tetraparvoviruses.

The alignment, in FASTA format, includes two sequences reported to come from endogenous viruses: that of *Indri indri* endogenous parvovirus and *Hyaena hyaena* endogenous parvovirus [1].

```
>ParvovirusB19/1-81
-----MDSYLT-T-PMPYHPVAVMQNLEEKMQYYLVKTYTSL
GKLAYNYPVLTMLGLAMS YKLGPRKVLLTVLQGFMTLGIANWLSWE-----
-----
>SimianParvo/1-81
-----MDLSQIS-EKLRQNTNPLMNLEEMTQSLLKKTYTKG
VELVLHYPLVTTLALVISFKLGIPKVWWT LRLGFMILGILNLLNWE-----
-----
>RhesusMacaque/1-79
-----MDLSLIS-ESLRQNMNPLMNLEEMTQSLVKRIYTKG
VDLVCNYP SVITLALVMSYRLGPRRVLWMLQLGVMISGILNLLS-----
-----
>PigtailedMacaque/1-79
-----MALSQIS-ENLR LSMRSLMNLEELTQSLLVKTYTMG
VQLVSHYPSLTILGLGILYKLG PQRTWWMLQLGFTITGILNLLS-----
-----
>IndriEndogenousParvovirus | RJWJ010000280/247-327
-----MNLTLVN-HLLNELNLTPPLLEQFLRDSGKTLTLYL
TEVVYTYPATATLVLATLWKTGLPLIWWT RPPGSTTSGTQKSKQKA-----
-----
>HyaenaParvovirus-like | PITC01058066.1_X protein Hyaena hyaena, whole
genome shotgun sequence 2000-3406/223->1405
MRGVIISQFQSTLLNQSLLLISAPPRSQRLGLRWRLMSLAPPS-TWRET----ILCLEN SLQRAISE-----
---VSHFPAVTTLGLAILW IQGPRLLSWTQRLAYTILGIHNYKAWGWILTPTLLRPTKSCCQPRLEKRDFRP
PW-----
>SealParvo/1-103
-----MSLSQNR-SWQEI----LSPLESGLQTMVIE-----
---ALHYPALTTLGLAIAWKKGPQLLGWTLRLAYTILGIVSSRSWALIPHTSHPLTRNSSRQPPEKRVSRQ
QWSRLSLS
>ChipmunkParvo/1-65
-----MW-QLFLNQTE-TLLQN-----NLNEAWQSLITGAST--
-----YPLTGILALAILWK TAPLLTRLTQWL VFTTLGTLT-----
-----
>Human_parvovirus_4/1-83
-----MSESTLLVT-AAMKE---LIPLLNQTL SG-LPEA----
---GLLSPGIIMLVLVILWIVVPLRDQW MRQQNIMMNGTQR-LSMGTS LIYMTALTD-----
-----
>PorcineHokoHK7/1-84
-----MEPIATVA-TVIKE---LVPLVNQTLNA-LPSAP---
---ESWSPGTIMLVLVILWIMLPLRDQW MKQRNTTMNGTMMKCFAMVICH TSMGEGLIG-----
-----
>YakHoko/1-84
-----MLDQATTVA-TVIRE---LVPLVNQTL SA-LPTA----
---DSWSPATIMLVLVILWIMVPLRDQW MRQRSTTMNGTMMKCCLMATCHMFMVTGQID-----
-----
>OvineHoko1/1-85
-----MAMAEVSTLA-TVVQE---LLPLVNQTLND-LPKG----
---DFWSPVTIMLVLVILWIMVPLRDQW TRQRNIMTNGTMMKCYHMGMSHIYTD TGLTD-----
-----
>DeerTetraparvo | NC_031670_1_cds_YP_009315887_1_2/1-/1-84
-----MESVVTVA-TVLEE---LIPLVNQTLNA-LPKG----
```

```

---DFLSPVIVMLVLVILWIMAPLRAQWTRPLNTMTNGTQICYLMGTYHTSMVMEQTDL-----
-----
>RodentTetraparvo/1-85
-----MADLLLPLS-SIMRE---LLPALNHTATE-LAREG---
---VSWSPGTIMLVILWIMAPLRDRLMRPQRDTMNGTMRCSSMGTCPICMAMGRMA-----
-----
>EidolonParvo1/1-77
-----MEEPLLNSALTSLKV---LAPLVNQTIHS-LAEV---
---VSLSPATIMWVLVILWIMAPLRAQWMRQLAIMTEGMTKCFPTGTSPT-----
-----
>OpposumTetraparvo/1-76
-----MEVSEGLQLLR-TIAQ-----ATANQTAEAILPGA---
---GSLSPAIIMLVILWIMAPLRAQWMRRLNAMTNGMMSCWVMGMCHT-----
-----

```

## References

1. Campbell MA, Loncar S, Kotin RM, Gifford RJ. Comparative analysis reveals the long-term coevolutionary history of parvoviruses and vertebrates. Quental TB, editor. PLOS Biol. 2022;20: e3001867. doi:10.1371/journal.pbio.3001867